



PDF hosted at the Radboud Repository of the Radboud University Nijmegen

This full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/15622>

Please be advised that this information was generated on 2014-11-12 and may be subject to change.

Segmental Phonology and the Perception of Syntactic Structure

DONIA R. SCOTT

University of Sussex

AND

ANNE CUTLER

MRC Applied Psychology Unit, Cambridge

Recent research in speech production has shown that syntactic structure is reflected in segmental phonology—the application of certain phonological rules of English (e.g., palatalization and alveolar flapping) is inhibited across phrase boundaries. We examined whether such segmental effects can be used in speech perception as cues to syntactic structure, and the relation between the use of these segmental features as syntactic markers in production and perception. Speakers of American English (a dialect in which the above segmental effects occur) could indeed use the segmental cues in syntax perception; speakers of British English (in which the effects do not occur) were unable to make use of them, while speakers of British English who were long-term residents of the United States showed intermediate performance.

Anyone who has had to make the separate meanings of an ambiguous sentence clear to listeners will know that often the most efficient method is just to say the sentence in different ways. Of course, some sentences will be more difficult to disambiguate in this manner than others; in general, surface structure ambiguities (e.g., “The old men and women stayed at home”) prove easier than deep structure ambiguities (e.g., “Flying planes can be

dangerous”) (Lehiste, 1973). But the fact that listeners can successfully identify the intended meaning even in the absence of a disambiguating context suggests that speakers can exploit acoustic features to highlight the distinction that is to be conveyed to the listener.

In recent years, a number of studies on English speech perception have investigated just what these acoustic correlates of syntactic structure are, and how useful they are to the listener. To date, such work has been almost exclusively concerned with the relationship between prosodic aspects of the speech wave (variations in fundamental frequency contour, duration, and amplitude) and the presence of major syntactic boundaries. This concern has to a large extent been motivated by findings which show that the prosody of an utterance can provide listeners with information about its surface structure; when listeners are presented with speech which has been modified (e.g., peak clipped, spectrally inverted, or band passed) in such a way that all seg-

This research was supported by a grant from the Science and Engineering Research Council to the University of Sussex. The authors are extremely grateful to Mark Pearson and John Local for the many hours they spent transcribing the production data; to Stephen Monsell for kindly providing space and equipment at the University of Chicago and access to the Chicago British community; to Steve Spearman for recording the stimuli; to Gill Brown and John Local for discussion of the distribution of palatalization and flapping in British English; and to Ilse Lehiste and Dennis Norris for helpful comments on the paper. Reprint requests should be addressed to the first author at the Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, United Kingdom.

mental information is destroyed but the prosodic pattern is retained, they are still able to identify the location of syntactic boundaries (Blessner, 1969; Kozhevnikov & Chistovich, 1965; Martin, 1972). The same is true for hummed or reiterant (where each syllable is replaced with "ma") versions of a sentence (Liberman & Streeter, 1978; Svensson, 1974).

Production studies have demonstrated that major syntactic boundaries are often accompanied by one or more of a number of prosodic features:

(a) a fall-rise in the fundamental frequency contour (Cooper & Sorensen, 1977; Lea, 1972; 't Hart & Cohen, 1973),

(b) a pause (Boomer, 1965; Goldman-Eisler, 1972; Grosjean & Deschamps, 1975),

(c) lengthening of the final stressed syllable of a phrase (Cooper, Paccia and Lapointe, 1978; Klatt, 1975; Lindblom & Rapp, 1973; Scott, 1982),

(d) lengthening of the foot (interstress interval) which contains the phrase boundary (Lehiste, Olive, & Streeter, 1976; Scott, 1982), and

(e) changes in amplitude (Streeter, 1978).

The first four of these features, at least, can be used by listeners as perceptual cues to the location of a major syntactic boundary (Collier and 't Hart, 1975; Lehiste et al., 1976; Scott, 1982; Streeter, 1978).

However, prosodic cues can be useful for disambiguation only if the speaker provides them, and the presence of prosodic markers of major syntactic boundaries in conversational speech will depend crucially on whether the speaker is aware of the ambiguity (Lehiste, 1973). But, even in the absence of prosodic cues, there may be other acoustic characteristics of the utterance which could be useful for disambiguation. For example, Cooper and his colleagues (Cooper and Paccia-Cooper, 1980; Egido and Cooper, 1980) have established that the syntactic structure of an utterance can ex-

ercise effects on *segmental* phenomena, namely, the application of certain phonological assimilation and elision rules. Many such rules can apply across a word boundary, but only if the word boundary is not also a major syntactic boundary. In this paper we attempt to assess whether such segmental effects can, like prosody, be used in perception, that is, can provide listeners with cues to the location of a phrase boundary. In particular, we examine the possible perceptual role of two rules studied by Cooper, namely, (a) flapping¹ and (b) palatalization of intervocalic alveolar stop consonants.

The phonological rule of flapping describes the change of intervocalic /t/'s and /d/'s to a flap ([ɾ]). In the production of a flap, the tongue tip moves rapidly upward and then downward, making brief contact with the alveolar ridge. The application of the flapping rule gives, for example, homophonous readings of the words *liter* and *leader* ([lɪɾ]). (Although it is sometimes claimed that vowel duration disambiguates flapped /t/'s from flapped /d/'s, Zue & Laffiere, 1979, reviewing their own and others' studies of medial flaps, conclude that /t/ and /d/ flap minimal pairs are in general perceived as homophonous). The flapping rule will also apply across a word boundary; thus the sequence /mɛt æn/ (met Ann) will be realized as [mɛɾ æn]. The rule of palatalization applies to intervocalic /t/'s and /d/'s in the environment of a following /j/; here the /t/ will be realized as [tʃ] and the /d/ as [dʒ]. Like flapping, palatalization will also occur across word boundaries, giving, for example, [dɪdʒ(j)u] for /dɪd ju/ (did you) and [mɛtʃ(j)u] for /mɛt ju/ (met you). The application of both rules, however, is inhibited when the word boundary between the stop and following segment is also a major syntactic boundary (Cooper et al., 1978; Egido & Cooper, 1980); Cooper

¹ In British phonetic terminology *flapping* is also referred to as *tapping* (see Wells, 1982, p. 249).

and his colleagues found that the /t/ or /d/ in the italicized portions of sentences 1 and 2 will be flapped in the (b) version of sentence 1 and palatalized in the (b) version of sentence 2 but not in the (a) versions, in which a major syntactic boundary comes between the stop and following segment.

- (1a) For those of you who'd like to *eat, early* lunch will be served.
- (b) For those of you who'd like to *eat early,* lunch will be served.
- (2a) We didn't break the *code yet* we intend to break it soon.
- (b) We didn't break the *code yet* but we intend to break it soon.

Because both the flapping rule and the palatalization rule are applied only optionally, the perceptual cues which they offer are somewhat indirect. Flapping never occurs immediately before a major syntactic boundary, but it does not have to occur in other environments either (Egido & Cooper, 1980); therefore, the *presence* of flapping should be a strong perceptual cue to the *absence* of a major syntactic boundary. The presence or absence of palatalization should be a somewhat less reliable syntactic marker in perception, since both forms are produced in both boundary and no-boundary conditions; but because palatalization occurs with significantly less likelihood when the boundary is present (Cooper & Paccia-Cooper, 1980) the *presence* of palatalization should also suggest that *no* major syntactic boundary is present. The present study examined whether listeners can indeed use the occurrence of flapping or palatalization as a cue to syntactic structure.

There is one further noteworthy feature of the rules of flapping and palatalization: their use is dialect specific. The distributional characteristics which Cooper and his colleagues described are typical of American English, but they do not correctly describe many other dialects of English: in particular, they do not apply to British English. In British English, flaps are virtually

unknown.² Palatalization occurs, but it occurs less frequently across a word boundary in British than in American English, and there is no evidence that its occurrence in British English is sensitive to the presence of a syntactic boundary.

This distributional pattern makes it possible to test a subsidiary hypothesis, namely, that the ability to see segmental phenomena of this type as cues to syntactic structure may be dependent on syntactically governed occurrence of the rules in question in one's own speech community. That is to say, given that only American speakers use flapping and palatalization as syntactically disambiguating cues, will only American listeners be able to derive syntactic information from such cues, or will the cues also be effective for British listeners (i.e., speakers who themselves do not mark syntactic structure in this way)? To test this hypothesis, we measured the sensitivity to syntactic patterning of these cues in groups of both American and British listeners.

Since the British neither produce such syntactic cues themselves nor hear them from their compatriots, we predict that although the cues will indeed be effective for the American listeners, the British listeners will not be able to make use of them. If this is indeed the case, further investigations will be necessary to identify the specific reason for the British listeners' inability to use such information. Two hypotheses present themselves: (a) the cues cannot be used because the British speakers do not themselves produce them (the Production Hypothesis) and (b) they cannot be used because the British speakers have had inadequate previous perceptual experience

² There are linguistic environments in British English where it can be appropriate to flap, but these environments are far more restricted than in American English. Flapping, when it occurs in British English, is most likely in monosyllabic words or short, rhythmically coherent (perhaps idiomatic) utterances, and only ever when the stop is preceded by a short vowel. Most British speakers, however, never flap /t/'s or /d/'s.

with them (the Perception Hypothesis). Although we may assume that all British listeners hear a fair amount of American speech, in the cinema and on television, if not in person, it may well be that a critical threshold of exposure has not been reached. Accordingly we included in our experiment a third group of listeners, namely, native speakers of British English who were long-term residents of America. If this group and the American group should prove able to use the segmental effects as cues to syntactic structure, while the British-resident British speakers could not, yet neither group of British speakers were to produce the effects, it would be appropriate to conclude that a critical degree of exposure to such effects is a sufficient prerequisite for their use as aids to parsing. On the other hand, if the American speakers should make perceptual use of the segmental information while neither of the British groups do, the implication would be that productive use of the discriminations in question is necessary for their employment in perception.

MATERIALS

The materials for this experiment consisted of the 12 syntactically ambiguous sentences shown in Table 1. The sentences were constructed in such a way that each sentence contained either a flapping environment or a palatalization environment at a potential phrase boundary (italic in the table). Half of the sentences were of the flapping type and the other half were of the palatalization type.

Two versions of each of the 12 sentences (one of each interpretation) were read by a male native speaker of American English. He was instructed to flap or palatalize the relevant /t/ or /d/ in readings where there was not a phrase boundary between the two critical words, but to release the stop when the phrase boundary was present. He was allowed time to practice and was also instructed to attempt to avoid pausing at the phrase boundary or producing contrastive

TABLE 1
TWELVE SYNTACTICALLY AMBIGUOUS SENTENCES
USED IN THE EXPERIMENT

Flapping sentences	
1.	The last time we <i>met Ann</i> was horrible.
2.	That time we <i>forgot Andrew</i> was really embarrassing.
3.	Each time we <i>visit India</i> is in the middle of the dry season.
4.	If you want to <i>eat early</i> lunch will be served.
5.	The costume must be <i>light orange</i> and made of natural fibres.
6.	Please buy some <i>tart apples</i> and a loaf of bread.
Palatalization sentences	
7.	I can see the people who <i>heard you</i> were here.
8.	I know the policeman who <i>suspected you</i> did it.
9.	All the books we <i>bought you</i> had designs on.
10.	John <i>believed you</i> like me.
11.	Mary <i>remembered you</i> like lightning.
12.	The man that <i>left you</i> wouldn't want to know.

Note. Six sentences contain a flapping environment, and six a palatalization environment at a potential phrase boundary. These sequences are italicized.

stress or intonational cues. His productions (several tokens of each version of each sentence) were recorded and later judged by the two authors according to these criteria.

Of each of the 24 resulting sentences, we chose the best token for the experiment. These were low-pass filtered at 4.5 kHz, digitized at a sampling rate of 10 kHz, and stored on a Vax computer. Two new, spliced versions were created for each of the chosen productions.

(1) The "two-word" version: Here the two words of the relevant flapping or palatalization environment were replaced by the same two words from the alternative reading of the same sentence.

(2) The "consonant" version: This was another hybrid sentence, where only the critical consonant of the flapping or palatalization environment was cross-spliced from the alternative reading.

The spliced sentences were prepared with the aid of the ILS speech package. Two-word splices were made from that point in the waveform where none of the word preceding the flapping or palatalization environment was audible, to that point

where none of the word following the same environment was audible. Consonant splices were made from the middle of the preceding vowel to the middle of the following vowel, with the middle being determined by counting pitch periods within each vowel.

All six versions of each sentence were then set to a "neutral" intonation contour which was the arithmetic mean of the pitch contours of the two original readings (measured over individual words). Original pitch readings were obtained via the LPC analysis facility of ILS and the new intonation contours were overlaid using an interactive pitch interpolation routine written by C. J. Darwin.

These manipulations enabled us to control for the presence of possible disambiguating cues in the original utterances, which could have enabled listeners to distinguish between phrase boundary and no-phrase boundary readings. (a) Intonational cues were removed since all final versions of a sentence now had the same intonation contour. (b) Timing cues were removed by the splicing manipulation in two ways (i) in the spliced versions timing cues would suggest one reading, while the critical segment would suggest the other, and (ii) in any case disruption of the smooth timing pattern of the utterance by the cross splicing would result in the normally efficient processing of rhythmic cues to syntax (Lehiste, 1977; Scott, 1982) being severely disrupted (Martin, 1979; Meltzer, Martin, Mills, Imhoff, & Zohar, 1976). This disruption applies equally to phrase boundary readings spliced into a no-phrase boundary context, and vice versa.

The experimental materials thus consisted of six versions of each of the 12 ambiguous sentences, each having the same intonation contour:

- (1) original phrase boundary reading,
- (2) phrase boundary context with no-phrase boundary two-word environment,
- (3) phrase boundary context with no-phrase boundary consonant,

- (4) original no-phrase boundary reading,
- (5) no-phrase boundary context with phrase boundary two-word environment, and

- (6) no-phrase boundary context with phrase boundary consonant.

Figure 1 shows an example of the six versions of the sentence "The last time we met Ann was horrible."

SUBJECTS

Three groups of subjects participated in the production and perception tests. One group (the British group) consisted of 11 native speakers of standard British English. A second group (the American group) consisted of 11 native American speakers who had been in Britain for less than 2 months at the time of the experiment (mean duration of stay was 3 weeks). A further 11 native British and 11 native American speakers participated in the plausibility ratings task. These groups were tested at Sussex University. The third production/perception group (the Chicago-British group) consisted of 14 Britons living in the United States (mean duration of stay in the United States was 4½ years). All were residents of the Hyde Park area of Chicago, Illinois, and were tested at the University of Chicago. All subjects were paid for their participation in the experiment.

PROCEDURE

The experiment consisted of two parts: a production test and a perception test. Before the subjects started the experiment, the ambiguous nature of each of the 12 sentences was explained to them.

Production Test

Subjects were given a list of sentences which consisted of each of the 12 sentences set in two different disambiguating context frames, one for each interpretation of the sentence. For example, the sentence "The last time we met Ann was horrible" was set in the two following contexts:

ORIGINAL PHRASE BOUNDARY (PB) READING	<u>The last time we</u> [mɛtæn] <u>was horrible</u>
PB READING WITH NO-PB 2-WORD ENVIRONMENT	<u>The last time we</u> [mɛɾæn] <u>was horrible</u>
PB READING WITH NO-PB CONSONANT	<u>The last time we</u> [mɛɾæn] <u>was horrible</u>
ORIGINAL NO- PHRASE BOUNDARY (NO-PB) READING	The last time we [mɛɾæn] was horrible
NO-PB READING WITH PB 2-WORD ENVIRONMENT	The last time we <u>[mɛtæn]</u> was horrible
NO-PB READING WITH PB CONSONANT	The last time we [mɛtæn] was horrible

FIG. 1. The six versions, two original and four spliced, of the flapping sentence: "The last time we met Ann was horrible." The flapping environment is shown in phonetic symbols. The underlined letters and symbols represent segments derived from the speaker's original phrase boundary reading ("Ann was horrible") and the non-underlined letters and symbols represent segments derived from the speaker's no-phrase boundary reading ("the occasion was horrible").

The first time we met her was a very enjoyable occasion but the last time we met Ann, (it) was horrible.

I don't know what I've done to offend her, but the last time we met, Ann was horrible (to me).

Each subject was then asked to read each sentence on the list (including the context but not the words in brackets) in what they considered to be a "natural" manner. Subjects were allowed to set their own pace and to repeat any reading which they were not satisfied with. Each subject's readings were recorded. The relevant flapping or palatalization environment of each production was later transcribed by one of the authors and by two colleagues trained in phonetic analysis.

Perception Test

Immediately after the recording session, subjects listened to a tape which contained three occurrences of each of the six prepared versions of each of the 12 sentences.

This was preceded by a trial session of 24 of the possible 72 utterances (two occurrences of each of the 12 sentences) in random order. The tape was divided into three blocks; each block contained each version of each sentence, presented in a pseudo-random order.³ There was a gap of 5 seconds between utterances and 2 minutes between blocks.

Subjects listened to the tape via headphones in a sound-proof cubicle. They were given an answer sheet with the alternative readings for each sentence and were instructed to mark which one was best expressed by the utterance heard.

Sentence Plausibility

Since it was possible that alternative readings of the sentences were not equally likely for the different groups of subjects, a further test was devised in which we col-

³ No sentence occurred twice in succession.

lected ratings of the meaning of the written text of each sentence on a scale from 1 to 5. A rating of 1 meant that the interpretation of the sentence in which no phrase boundary is present between the two critical words was much more likely; a rating of 5 meant that the interpretation with a phrase boundary between the two critical words was much more likely; a rating of 3 meant that both readings were equally plausible.

Two new groups of subjects provided rating controls for British and American subjects: 11 native British speakers living in Britain and 11 native American speakers who, as before, had been in Britain for less than 2 months. Unfortunately, we were unable to locate a sufficient number of British speakers in Chicago in the time available to provide a rating control group for our Chicago-British speaker/listeners. This group of subjects therefore acted as their own rating controls, and were given the rating test after they had completed the perception test.

RESULTS

Plausibility Judgments

A criterion was set whereby those sentences which received a mean plausibility rating within the range of 1.5 to 4.5 were considered to be ambiguous, with sentences falling outside this range being heavily biased toward one or the other interpretation. Mean rating judgments for each sentence are given in Table 2. As can be seen in this table, not all sentences were found to be ambiguous, and sentences which are ambiguous for one group are not necessarily so for all three groups.

Two of the flapping sentences (sentence 2 and sentence 5) were judged to be unambiguous by the British group (with the no-phrase boundary reading being the most plausible one). Sentence 5 was also unambiguous (in the same direction) for the American group, which also judged sentence 6 as being strongly biased in favor of

TABLE 2
MEAN RATING JUDGMENTS FOR THE 12 AMBIGUOUS
SENTENCES LISTED IN TABLE 1

	Americans (<i>n</i> = 11)	British (<i>n</i> = 11)	Chicago-British (<i>n</i> = 14)
1.	2.1	2.9	2.8
2.	1.5	1.2*	2.4
3.	2.1	3.4	2.7
4.	1.7	2.0	2.1
5.	1.4*	1.4*	2.0
6.	1.4*	2.1	2.2
7.	4.0	4.3	2.6
8.	4.8*	4.5	3.4
9.	2.7	2.4	2.8
10.	4.1	3.9	2.2
11.	4.0	2.5	2.2
12.	2.8	2.8	3.2

Note. A rating of 1 indicates that the sentence is heavily biased toward the no-phrase boundary reading; a rating of 5 indicates that it is heavily biased toward the phrase boundary reading. Sentences which fall outside of our criterion of "ambiguous" are marked with an asterisk.

the no-phrase boundary reading. For the Chicago-British group, all six flapping sentences were rated as ambiguous. It is difficult to say whether the ratings received for this group provided adequate insight into the plausibility of the respective readings, since these raters had previously been subject to repeated exposure to the same sentences in a situation where they were being forced to consider the two possible interpretations of each sentence. The Chicago-British group also judged all six of the palatalization sentences as being ambiguous. Only one of the palatalization sentences was judged to be unambiguous; for the American group, sentence 8 was rated as being strongly biased toward the phrase boundary reading.

Perception

For each group, responses to only those sentences which were judged to be ambiguous in the rating test were submitted to statistical analysis. Subjects' responses were scored in terms of the number of times (out of a total of three) that the reading containing a phrase boundary between the two

TABLE 3
MEAN PHRASE BOUNDARY JUDGMENTS FOR
PALATALIZATION SENTENCES

	Original	Two- word	Consonant
Americans			
Not palatalized	2.09	1.87	1.64
Palatalized	1.27	1.33	1.14
British			
Not palatalized	1.35	1.41	1.33
Palatalized	1.11	1.12	1.19
Chicago-British			
Not palatalized	1.66	1.76	1.74
Palatalized	1.76	1.41	1.27

critical words was chosen for each stimulus item. These scores were subjected to separate analyses of variance with subjects and with items as random factors, in order to allow calculation of $\min F'$, which indicates whether results will generalize both to further subjects and to other sentences. Since the results of the rating test, reported above, had led to the exclusion of certain sentences from the calculation of the results for two of the subject groups, the three sets of results were therefore not exactly comparable, and were hence analyzed separately.

Palatalization. The mean numbers of phrase boundary readings chosen for each group and each stimulus condition are presented in Table 3. The results were quite different for the three groups of subjects. For the American subjects, sentences with palatalization attracted significantly fewer phrase boundary judgments than sentences without palatalization ($\min F'(1,6) = 7.19$, $p < .05$). The effect of context (original, two-word, or consonant only), although significant in the analysis by subjects ($F(2,20) = 7.88$, $p < .01$), did not reach significance on the combined analysis ($\min F'(1,11) = 1.11$), and did not interact with the palatalization effect ($\min F' < 1$). Thus the effect of palatalization was independent of whether the palatalized consonant was accompanied by all, a little, or none of its original context.

TABLE 4
MEAN PHRASE BOUNDARY JUDGMENTS FOR
FLAPPING SENTENCES

	Original	Two- word	Consonant
Americans			
Not flapped	2.23	2.11	2.00
Flapped	.66	1.14	1.46
British			
Not flapped	1.50	1.25	.98
Flapped	.73	.66	.94
Chicago-British			
Not flapped	1.70	1.68	1.69
Flapped	1.00	.69	1.01

For both the British and the Chicago-British subjects, however, no effect at all reached the set level of significance ($\min F' < 1$ on all analyses for the British and on both main effects for the Chicago-British; interaction of context and palatalization for the Chicago-British $\min F'(1,16) = 2.67$). In particular, the main effect of palatalization found with the American subjects was not displayed by either of the other two subject groups.

Flapping. The mean numbers of phrase boundary judgments for each group and each stimulus condition are displayed in Table 4. Again, the three groups exhibit different response patterns. For the American subjects, sentences with flaps attracted significantly fewer phrase boundary judgments than sentences without flaps ($\min F'(1,4) = 9.08$, $p < .05$). The same effect was shown by the Chicago British subjects ($\min F'(1,8) = 7.64$, $p < .05$). For the British subjects, however, although the flapping effect did reach significance in the analysis by subjects ($F(1,10) = 17.92$, $p < .01$), it was not significant in the combined analysis ($\min F'(1,4) = 3.41$).

Neither the effect of context itself nor the interaction of the flapping effect with type of context reached significance for any of the three groups ($\min F' < 1$ on both analyses for both British groups; for the Americans $\min F'(2,9) = 1.58$ for context effect, $\min F'(2,10) = 2.18$ for interaction). In

other words, the effect of flapping, in the two groups that showed it, was independent of the amount of original context also present.

Production

The taped readings of each subject for each version of each of the 12 ambiguous sentences were transcribed by one of the authors and by two colleagues trained in phonetic analysis. As we were only interested in the number of times the subjects flapped or palatalized the /t/ or /d/ at the locations italicized in Table 1, only these portions of each utterance were transcribed.

Transcription took place in two stages. First, each transcriber independently judged all of the utterances. Any token which was judged to be flapped or palatalized with less than 100% agreement was submitted to a "second pass" analysis, in which the transcribers listened to the items jointly and came to a mutually agreed decision about what the token was.⁴

Segments were marked as flaps only when there was a clearly perceptible flap. These sounds were extremely easy to identify, especially since all possible flaps were derived from the unvoiced /t/ and flaps are usually voiced.⁵ In the "first pass," there were only 2 out of a total of 432 tokens on which the transcribers disagreed.

Decisions about the presence or absence of palatalization were much more difficult to make than those about flapping. There were a large number of tokens where a decision on palatalization could be made only after listening to the utterance over and over again at half the normal playback speed. The presence of friction between the

stop and following segment was not always an adequate criterion for palatalization; there were a few cases where friction was present, but where the transcribers felt that the utterance was not palatalized. Only sounds which were good representatives of the initial phoneme in *judge* or *church* (depending on whether the underlying stop was a /d/ or /t/) were accepted as palatalized.

The percentages of palatalized and flapped tokens for the different syntactic conditions and for the three groups of subjects, averaged over subjects, are shown in Table 5.⁶

In the case of flapping, the results showed that none of the subjects flapped in any of the readings where a phrase boundary was meant to occur between the /t/ and following segment. The American group, however, flapped in over 45% of the readings where the phrase boundary was not present. Only one British subject flapped at all, and then only in one (no-phrase boundary) reading. The Chicago-British subjects flapped in just 5% of no-phrase boundary readings.

Unlike flapping, palatalization was found to occur in both syntactic conditions. Again, the British and Chicago-British groups gave a similar pattern of results and both groups differed from the Americans. The Americans palatalized more than twice as frequently when the phrase boundary was absent than when it was present. The British and Chicago-British, however, made much less distinction between syntactic structures. The Americans, then, appear to be marking the absence of a phrase boundary in production by flapping or palatalization, whereas the two British groups do not.

It is interesting to note that our American subjects' production performance closely parallels that reported in the production

⁴ Two transcribers, M.P. and D.S., went through their transcriptions together; final results of this session were then submitted to a further session with the third transcriber, J.L.

⁵ There were a few unvoiced flaps present in the data. This is not an unusual finding. Fox and Terbeek (1977) report that flaps (derived from either voiced or unvoiced alveolar stops) are sometimes unvoiced.

⁶ Since the three groups are not represented by equal numbers of subjects, the results are given in the form of percentages to aid comparison between groups.

TABLE 5
PERCENTAGE OF FLAPPED OR PALATALIZED
UTTERANCES (SUMMED OVER SPEAKERS) IN THE
PHRASE BOUNDARY (+PB) OR NO-PHRASE
BOUNDARY (−PB) READINGS FOR THE
AMERICANS (A), BRITISH (B), AND CHICAGO-
BRITISH (C)

	A	B	C
Flapping			
+ PB	0	0	0
− PB	45.5	1.5	4.8
Difference	45.5	1.5	4.8
Palatalization			
+ PB	19.7	25.8	25.0
− PB	43.9	34.8	33.3
Difference	24.2	9.0	8.3

studies which formed the stimulus for the present study. Egido and Cooper (1980) found that with ambiguous sentences of the type we used (in fact, our sentence 4 was taken from their experiment) 40% of (American) speakers produced flaps in the no-phrase boundary reading, but none at all in the phrase boundary reading. Cooper et al. (1978), testing a somewhat different syntactic effect, namely, whether the word boundary in question was a deletion site, found that roughly twice as many speakers palatalized across the minus-deletion (56%) as across the plus-deletion (32%) boundary.

The results given in Table 5, it should also be noted, do not do justice to the richness of the data. The phonological rules of flapping and palatalization can be blocked by factors other than the presence of a major syntactic boundary. For example, palatalization is much less likely to occur when there is emphatic stress on one of the two relevant words, especially when there is stress on the second (/j/) word (Cooper, Soares, Ham, & Damon, 1982). Similarly, it is generally accepted that the alveolar flapping rule applies only in post-stress position or where both words of the flapping environment bear equal stress (i.e., not if the (/j/) word has a higher stress value than the preceding /t/ or /d/ word) (Cooper et al., 1978; Lorge, 1967; Oshika, Zue, Weeks,

Neu, & Aurbach, 1975). The occurrence of palatalization is also affected by speech rate. Cooper et al. (1982) found that characteristically fast speakers tend to palatalize more often than slow speakers and that both groups palatalize more often when instructed to speak at a faster-than-normal rate and less often when speaking at a slower-than-normal rate. A similar effect of rate may operate on flapping which, by definition, requires a rapid articulatory gesture. The subjects in this experiment were only instructed to read the sentences "as though they were saying them naturally." No other constraints were put on their productions. There were a number of cases where speakers placed stress on one of the flapping or palatalization words. We also suspect that some speakers went into "elocutionary mode" during the recording session, speaking more slowly and clearly than they would in normal conversational speech. The figures given in Table 5 do not take into account the operation of the above factors which, although they should not affect the relative difference in the number of flapped or palatalized tokens across syntactic conditions, do affect the absolute number of flapped or palatalized tokens. The figures given also do not reflect the fact that some readings were felt by the transcribers to be clearly wrong (i.e., it was felt that the speaker's intended meaning was not the one which the context suggested). Since this impression was based entirely on the way in which the sentence was said, these utterances were not eliminated. Finally, these figures do not reflect any interspeaker or intersentence differences; nor do they tell us what speakers were producing on those occasions when they were neither flapping nor palatalizing the stop.

DISCUSSION

The major hypothesis which this experiment was designed to test has been unequivocally supported by the results; the occurrence of segmental variations which

in production are sensitive to syntactic structure is indeed informative to listeners engaged in the parsing of syntactically ambiguous utterances. The presence of a flapped consonant at a word boundary serves as an effective cue to the absence of a phrase boundary between the two words in question. Similarly, the presence of palatalization across a word boundary is used as a signal that no phrase boundary occurs at that point. Flapping and palatalization are presumably not the only segmental rules which carry syntactic information; our results suggest that any segmental effect with a syntactically systematic pattern of distribution will be as effective a disambiguating cue as previous research has shown prosodic effects to be.

However, not all speakers of English are equally able to use segmental cues in this manner. Simply speaking, those who produce the cues in their own speech can make best use of them in perception. Only our American subjects consistently produced both of the segmental effects in a syntactically differentiated manner *and* consistently used them in perception as cues to syntactic structure.

The British listeners, on the other hand, who (as predicted) did not produce either effect in a systematic way, also failed to be able to make any use of either effect as a cue to syntactic structure. Thus if one neither produces such effects in one's own speech, nor regularly hears them in the speech of others, one cannot derive disambiguating information from them.

The contrast between the British and the Chicago-British listener groups, it will be recalled, was specifically designed to provide further information in the event of just such a pattern of results being found across the American and British groups. We predicted that if production were the crucial prerequisite (the Production Hypothesis), then the Chicago-British, assuming they did not produce the effects, would also not show them in perception. On the other hand, if perceptual exposure were a suffi-

TABLE 6
RELATIONSHIP BETWEEN THE PRODUCTION AND PERCEPTION OF, AND EXPOSURE TO, FLAPPING AND PALATALIZATION AS MARKERS OF A PHRASE BOUNDARY FOR THE THREE GROUPS

	Exposure	Production	Perception
		Palatalization	
Americans	Yes	Yes	Yes
British	No	No	No
Chicago-British	Yes	No	No
		Flapping	
Americans	Yes	Yes	Yes
British	No	No	No
Chicago-British	Yes	No	Yes

cient prerequisite (the Perception Hypothesis), then the Chicago-British would show the perceptual effects.

Our results appear, however, not to distinguish unequivocally between the Production and Perception hypotheses. Table 6 shows the pattern of perceptual exposure, productive use and perceptual use of the groups. The Chicago-British group are exposed to both flapping and palatalization as syntactic markers in the speech of those around them. However, the production test shows that they use neither rule in the way that their American neighbors do. If the Production Hypothesis were correct, they should therefore use neither in perception; if the Perception Hypothesis were correct, they should use both. In fact, they appear to use one but not the other.

Superficially it would appear, then, that the Production Hypothesis holds true for palatalization, while the Perception Hypothesis is true of flapping. If this were indeed the case, it would be necessary to look for some fundamental difference between the two rules which could account for such a radical difference in the way they are handled by the human language processor. However, we shall claim that no such categorization of types of rules is in fact necessary. Instead, we shall argue that the whole pattern of our results can be (at least

cautiously) interpreted as providing support for the Production Hypothesis.

In the case of palatalization, the argument is clear. The British and Chicago-British groups were strikingly similar in the pattern of their productions, both palatalizing in 25–35% of phrase boundary *and* no-phrase boundary versions of the test sentences. The Americans on the other hand palatalized significantly more often in the no-phrase boundary condition than they did in the phrase boundary condition. That the Americans then showed the perceptual effect while neither the British nor the Chicago-British did is precisely the pattern of results predicted by the Production Hypothesis.

The flapping results are more complicated. The British did not flap at all in production, while the Americans did but only in no-phrase boundary contexts. The Chicago-British, like the British, virtually never produced flaps, and according to the Production Hypothesis they should therefore not have been able to use flaps as perceptual cues. However, their use of them, though not as successful as the Americans', was quite efficient. The key to this anomaly lies in the production data. Close inspection of the phonetic transcriptions of our subjects' utterances reveals that although a mere tally of the number of flaps gives almost identical results for British and Chicago-British speakers, the utterances which these two groups of speakers produced were in fact different in many respects. The British speakers aspirated the [t] sound in virtually all utterances, that is, produced nearly 100% unvoiced, audibly released stops irrespective of the presence or absence of a phrase boundary. The Chicago-British speakers, on the other hand, produced a much greater variety of sounds: a relatively large number of unexploded [t]'s (i.e., [t]'s produced with overlapping glottal stop), and in some cases even a definite [d], that is, a fully voiced alveolar stop. They also produced more flaps and glottal stops than the British did. Figure 2 shows spec-

trograms of the words "visit India" from the no-phrase boundary version of sentence 3 as produced by (a) a typical British speaker, using an exploded [t]; (b) a typical American speaker, using a flap; and (c–e), three Chicago-British speakers using in turn an unexploded [t], a [d], and a double glottal stop.

This pattern of results is consistent with the hypothesis that the Chicago-British speakers are actually in the process of acquiring the American-English phonological rule which turns intervocalic /t/ to a flap. 18.5% of the consonants they produced in the critical environments were not the aspirated, unvoiced stop which British English prescribes for such positions. This may not seem a large proportion—but recall that the circumstances in which these sentences were produced generally elicit slow, careful speech, and therefore tend to reduce the likelihood of the rule being applied. Even the American native speakers only produced a total of 25.8% responses which were not aspirated and unvoiced. Furthermore, informal inspection of the rest of the taped material produced by our speakers revealed numbers of corroborative examples. The disambiguating contexts for sentence (1), for instance, were:

+PB: I don't know what I've done to offend her, but the last time we met, Ann was horrible.

–PB: The first time we met her was a very enjoyable occasion, but the last time we met Ann was horrible.

Several of the Chicago speakers produced [d] instead of [t] in "what I've" or in "met 'er"—which, of course, are also appropriate contexts for application of the flapping rule. The American speakers tended to produce flaps here too; the British, however, uniformly produced exploded [t]'s.

Admittedly, the Chicago-British group's acquisition of the correct form of the rule would seem to be less than perfectly successful so far. Their productions might be

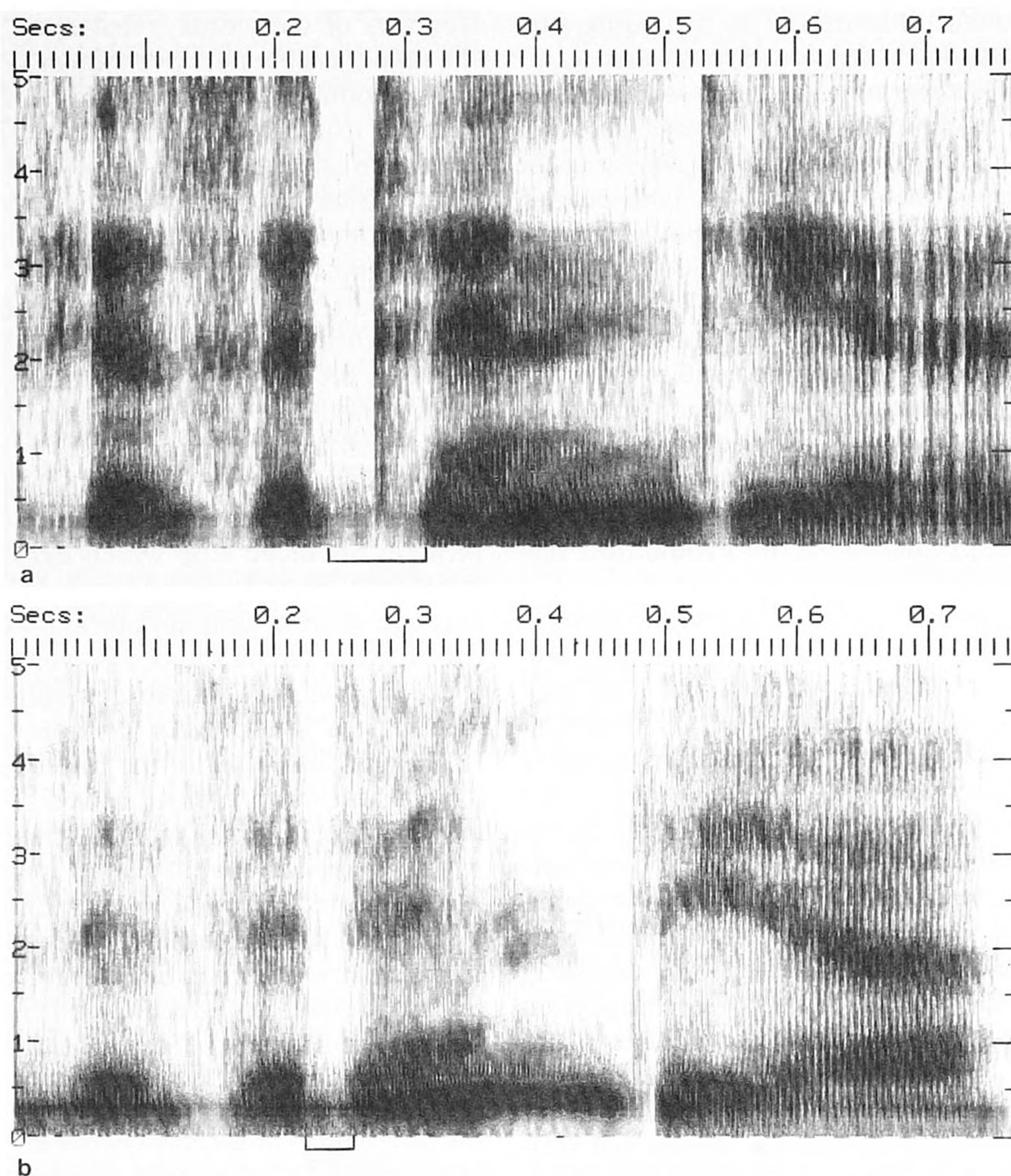
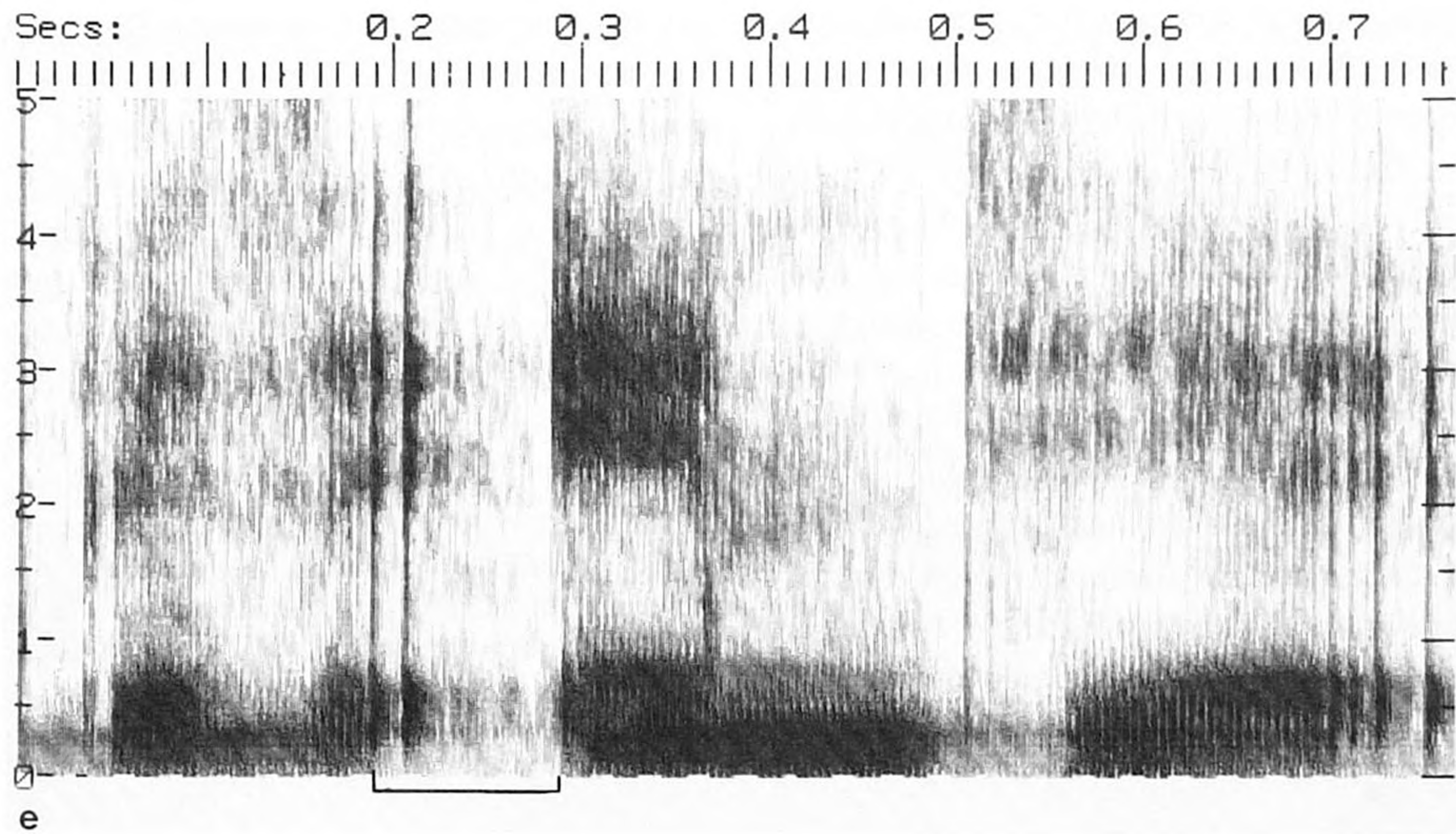
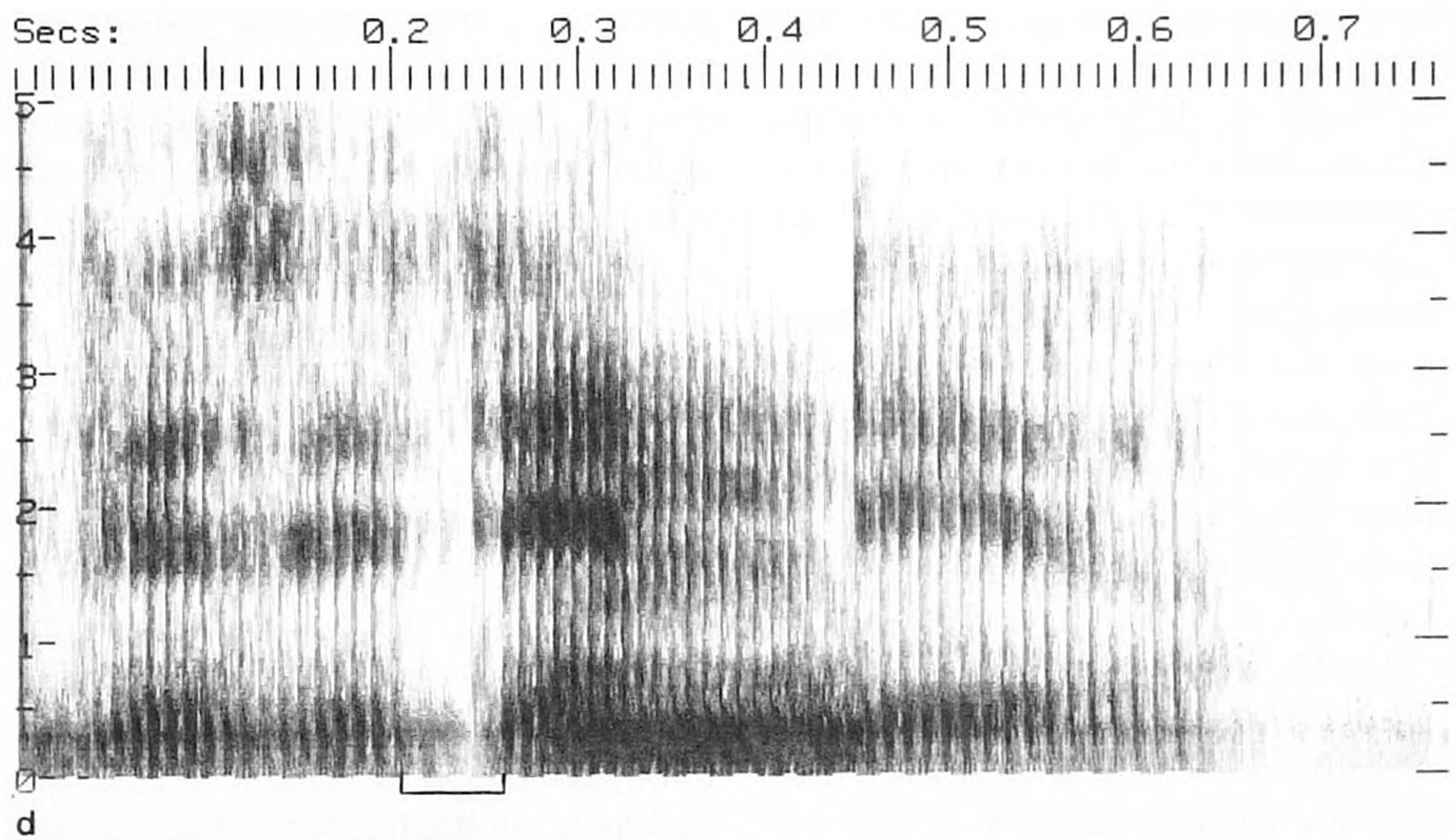
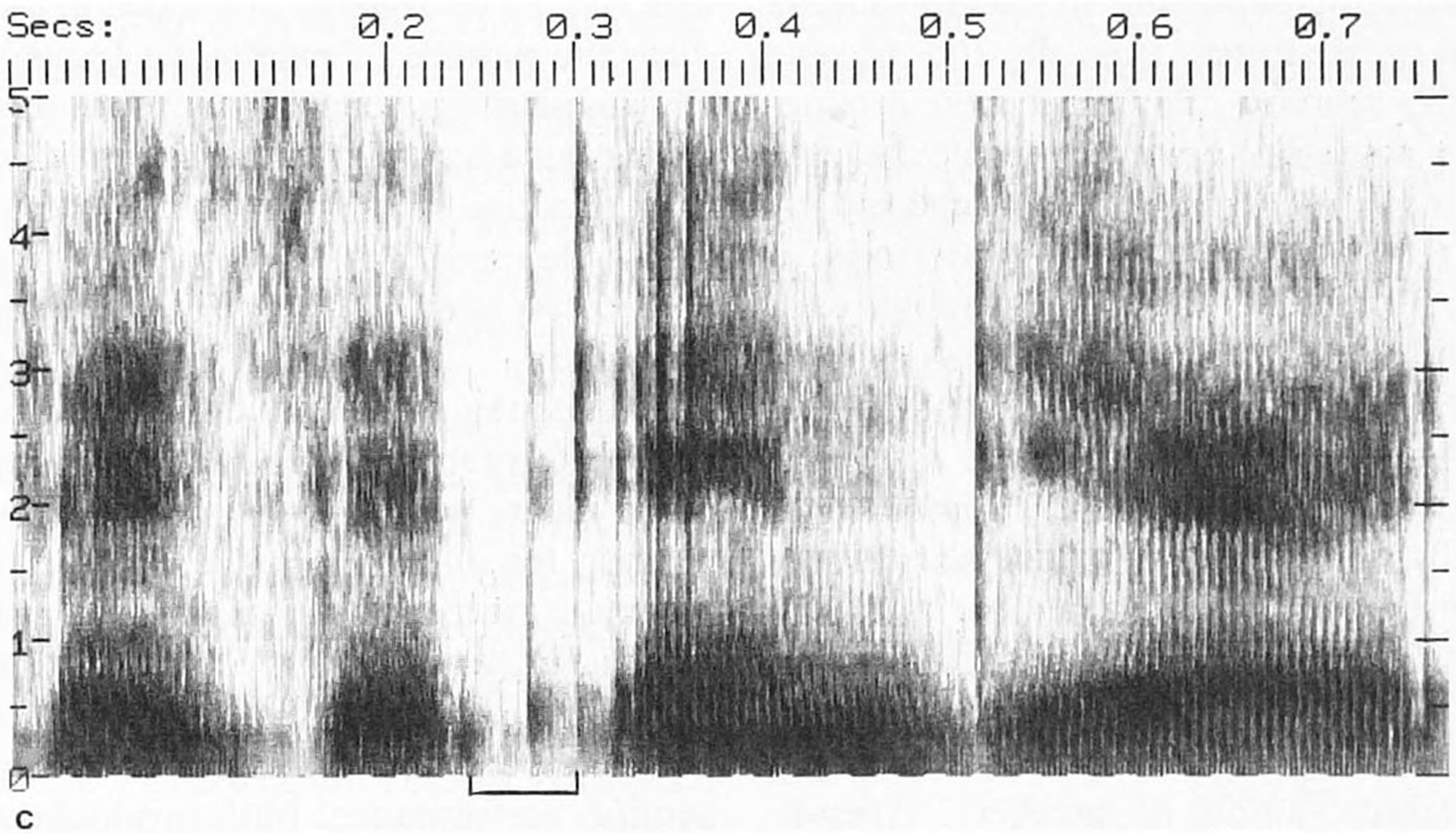


FIG. 2. Spectrograms of the words "visit India" from five productions of the no-phrase boundary version of sentence 3. In each case the black line represents the critical consonant portion, that is, the portion from the offset of the preceding vowel to the onset of the following vowel. In (a), spoken by a British speaker, the consonant is an aspirated [t]. The aspiration can be seen from approximately 0.28 to approximately 0.32 seconds on the time scale. In (b), spoken by an American, there is a flap. (c–e) are spoken by Chicago-British speakers; (c) contains an unexploded [t] (comparison with (a) shows the aspiration portion in (a) to be missing; the vowel is initiated with a glottal release at approximately 0.28 seconds on the time scale); (d) contains a [d], and (e) a double glottal stop. In both (d) and (e), as in (b), voicing continues through the consonant (i.e., the darker bands, representing the formants, can be seen in the consonant portion of the spectrogram).

characterized as "flapping with a British accent." The sounds they produced were in fact very rarely true flaps. However, they have correctly assumed that the desired target sound is unaspirated and voiced; and the production of a [d] is not really all that

incorrect either, since voicing does continue through a flap—to an unaccustomed British ear, a flap sounds much more like a /d/ than a /t/. Furthermore, there was some indication that the Chicago speakers were producing systematically more of these at-



tempted-flap responses in the no-phrase boundary readings than in the phrase boundary readings: 26.2% of their productions in no-phrase boundary contexts were classified as other than audibly released [t], but only 10.7% of their productions in phrase boundary contexts (although with the small numbers involved, these proportions are not significantly different). There was no indication that length of residence in America (our group varied from 1 to 20 years) was correlated with either production or perception performance; but individual differences in rate of accent acquisition are very great, so that such a question could only be answered by studying a much larger sample of speakers. Overall, our results certainly seem to indicate that the Chicago-British speakers are undergoing a change in their speech patterns, such that both their production and perception performance falls between that of the British and American groups.

The flapping results, therefore, also tend to support the Production Hypothesis rather than the Perception Hypothesis. These two hypotheses, as we formulated them above, were intended to account for a predicted difference in perception performance between the British and American groups, which in the event we indeed found: Americans use the phonological effects we studied in both production and perception; the British use them in neither. According to the Perception Hypothesis, the British listeners would be unable to use flapping and palatalization as syntactic disambiguators because of inadequate previous perceptual exposure to them. According to the Production Hypothesis they would be unable to use them in perception because they did not themselves use them in production. The Perception Hypothesis is clearly not supported by our results. The Chicago-British group had adequate perceptual exposure to the use of palatalization as a syntactic disambiguator but failed to be able to make perceptual use of it in this experiment. The Production Hypoth-

esis, we have argued, more correctly predicts the pattern of results we found—the Chicago-British group do not use palatalization as a syntactic marker, neither do they produce it as such; they make some use of flapping in perception, and they appear to be acquiring its use in production as well.

In claiming that production and perception performance run parallel, it should be made clear, we make no claim of logical priority for either language production or language perception in linguistic competence. We would prefer to view competence as a unitary phenomenon at a more abstract level than either production or perception performance; both production and perception performance would then be equally dependent upon competence. If a rule has been acquired, in other words, it can be applied equally well in production and in perception. The syntactic sensitivity of the flapping and palatalization rules, once acquired, can be displayed in speech and made use of in parsing with equal facility.

Finally, we consider why the Chicago-British should apparently be in the process of acquiring the flapping rule, along with its syntactic sensitivity, from their American environment, while remaining uninfluenced by American palatalization effects. The answer here lies in the characteristics of these speakers' dialectal base, that is, standard British English. The alveolar flap is not a feature of British English. Thus flapping is a very distinctive characteristic of American speech to British ears; because it is distinctive, it is readily available to be adopted by British speakers who are accommodating their speech to their new environment.⁷ (Of course, as our production results show, the successful acquisition of

⁷ There may, in addition, be a social penalty for British speakers using intervocalic [t] in the United States. Even more than the British accent in general, "the typical RP . . . voiceless alveolar plosive . . . is often perceived by Americans as artificial, prissy or effeminate" (Wells, 1982, p. 250).

such phonological features turns out to be not a simple matter).

Palatalization, on the other hand, *does* occur in British speech. Moreover, in British English the number of available environments for palatalization is in fact far greater than in American English. For example, the initial sounds of the words *duty* and *Tuesday* can be palatalized by British speakers, because there is a glide /j/ between the initial stop and the vowel; such environments have disappeared from American English. Thus palatalization is a far less selective effect in British than in American English. In order to adapt their use of the palatalization rule to American standards, British speakers would have to reduce the range of potential environments for its application.

Moreover palatalization in British English already carries nonphonological information; its occurrence is characteristic of nonstandard speech. Therefore its use tends to be, for native speakers of British English, a marker of social class. We would suggest that the Chicago-British are in the process of acquiring (a version of) flapping from their new speech environment because flapping is an entirely new linguistic device with respect to their dialectal base. They have failed to acquire American patterns of palatalization, however, because (a) palatalization is a more general effect in British English, so that acquisition of American palatalization patterns would necessitate a *reduction* in its use, which may be more difficult than simple addition of a phonological rule to ones repertoire and (b) the availability of palatalization as a carrier of nonphonological (syntactic) information has been preempted by its function as a carrier of a (nonphonological) sociological message.

Inability of speakers to acquire a new use for a contrast which already has an informative function in their dialect has been established with other phenomena. For example, Berinstein (1979) found that speakers of Mayan languages with fixed

final stress could learn to use durational variation as a position-independent cue to word stress (in the way that English speakers do), *but only* if their language was one which did not have phonemic vowel length variations—that is, only if their language had not already preempted durational variation as a cue to something other than stress.

APPENDIX PHONETIC SYMBOLS USED IN THE TEXT

Vowels	
ɛ	as in <i>met</i>
æ	as in <i>Anne</i>
ɪ	do in <i>did</i>
i	as in <i>eat</i>
u	as in <i>you</i>
Consonants	
j	as in <i>you</i>
d	as in <i>do</i>
t	as in <i>too</i>
l	as in <i>low</i>
r	as in <i>row</i>
m	as in <i>mow</i>
n	as in <i>no</i>
ʃ	as in <i>shoe</i>
ʒ	as in <i>rouge</i>
tʃ	as in <i>chew</i>
dʒ	as in <i>Joe</i>
ɾ	(flap—see text)

REFERENCES

- BERINSTEIN, A. E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics*, 47.
- BLESSER, B. A. (1969). *Perception of spectrally rotated speech*. Unpublished doctoral thesis, Massachusetts Institute of Technology.
- BOOMER, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech*, 8, 148–158.
- COLLIER, R., & 'T HART, J. (1975). The role of intonation in speech perception. In A. Cohen and S. G. Nooteboom (Eds.), *Structure and process in speech perception* (pp. 107–121). Heidelberg: Springer-Verlag.
- COOPER, W. E., EGIDO, C., & PACCIA, J. M. (1978). Grammatical control of a phonological rule: Palatalization. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 264–272.
- COOPER, W. E., PACCIA, J. M., & LAPOINTE, S. G.

- (1978). Hierarchical coding in speech timing. *Cognitive Psychology*, 10, 154–177.
- COOPER, W. E., & PACCIA-COOPER, J. M. (1980). *Syntax and speech*. Cambridge, MA: Harvard Univ. Press.
- COOPER, W. E., SOARES, C., HAM, A., & DAMON, K. (1982). Planning speech for execution at different tempos. *Journal of the Acoustical Society of America*, 72 (Suppl. 1), S64.
- COOPER, W. E., & SORENSEN, J. M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America*, 62, 683–692.
- EGIDO, C., & COOPER, W. E. (1980). Blocking of alveolar flapping in speech production: The role of syntactic boundaries and deletion sites. *Journal of Phonetics*, 8, 175–184.
- FOX, R. A., & TERBEEK, D. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics*, 5, 27–34.
- GOLDMAN-EISLER, F. (1972). Pauses, clauses, sentences. *Language and Speech*, 15, 103–113.
- GROSJEAN, F., & DESCHAMPS, A. (1975). Analyse contrastive des variables temporelles de l'Anglais et du Français: Vitesse de parole et variables composantes, phénomènes d'hesitation. *Phonetica*, 31, 144–184.
- KLATT, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, 3, 129–140.
- KOZHEVNIKOV, V. A., & CHISTOVICH, L. A. (1965). *Speech: Articulation and perception*. Washington, DC: U.S. Department of Commerce Translation, J.P.R.S., 30, 543.
- LEA, W. A. (1972). *Intonational cues to the constituent structure and phonemics of spoken English*. Unpublished doctoral thesis, Purdue University.
- LEHISTE, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 107–122.
- LEHISTE, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5, 253–263.
- LEHISTE, I., OLIVE, J. P., & STREETER, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America*, 60, 1199–1202.
- LIBERMAN, M., & STREETER, L. (1978). Use of nonsense-syllable mimicry in the study of perceptual phenomena. *Journal of the Acoustical Society of America*, 62, 231–233.
- LINDBLOM, B., & RAPP, K. (1973). Some temporal regularities of spoken Swedish. *Papers from the Institute of Linguistics, University of Stockholm*, 21.
- LORGE, B. (1967). Production and perception of initial and intervocalic /t/ and /d/ in English. *Haskins Laboratory, Status Report on Speech Research*, 9, 11–18.
- MARTIN, J. G. (1972). Rhythmical (hierarchical) vs. serial structure in speech and other behaviors. *Psychological Review*, 79, 487–509.
- MARTIN, J. G. (1979). Rhythmic and segmental perception are not independent. *Journal of the Acoustical Society of America*, 65, 1286–1297.
- MELTZER, R. H., MARTIN, G., MILLS, C. B., IMHOFF, D. L., & ZOHAR, D. (1976). Reaction time to temporally-displaced phoneme targets in continuous speech. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 277–290.
- OSHIKA, B. T., ZUE, V. W., WEEKS, R. V., NEU, H., & AURBACH, J. (1975). The role of phonological rules in speech understanding research. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 23, 104–112.
- SCOTT, D. R. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America*, 71, 996–1007.
- STREETER, L. A. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America*, 64, 1582–1592.
- SVENSSON, S. G. (1974). Prosody and grammar in speech perception. *Monographs from the Institute of Linguistics, University of Stockholm*, 2.
- 'T HART, J., & COHEN, A. (1973). Intonation by rule: A perceptual quest. *Journal of Phonetics*, 1, 309–327.
- WELLS, J. (1982). *Accents of English* (Vol. 1). Cambridge: Cambridge Univ. Press.
- ZUE, V. W., & LAFERRIERE, M. (1979). Acoustic study of medial /t,d/ in American English. *Journal of the Acoustical Society of America*, 66, 1039–1050.

(Received February 24, 1983)

(Revision received June 14, 1983)